改进粒子群优化的极限学习机软测量建模方法 *

盛晓晨 [a]，史旭东 [a]，熊伟丽 [a, b]

(江南大学 a.物联网工程学院，江苏 无锡 214122; b. 轻工过程先进控制教育部重点实验室，江苏 无锡 214122)

摘　要：工业过程常含有显著的非线性、时变等复杂特性，传统的极限学习机有时无法充分利用数据信息，故所建软测量模型预测性能较差。为了提高极限学习机的泛化能力和预测精度，提出一种改进粒子群优化的极限学习机软测量建模方法。首先，利用高斯函数正态分布的特点实现惯性权重的自适应更新，并线性变化学习因子，以提高粒子群优化算法的收敛速度和搜索性能；然后，将该算法用于优化极限学习机的惩罚系数和核宽，得到一组最优超参数；最后，将该方法应用于脱丁烷塔过程软测量建模中。仿真结果表明，优化后的极限学习机模型预测精度有明显的提高，验证了所提方法不仅是可行的，而且具有良好的预测精度和泛化性能。

关键词：软测量建模；极限学习机；粒子群优化算法；自适应权重

中图分类号：TP273　　doi: 10.19734/j.issn.1001-3695.2018.11.0863

Soft sensor modeling for extreme learning machine based on improved particle swarm optimization

Sheng Xiaochen[a], Shi Xudong[a], Xiong Weili[a, b]

(a. *School of Internet of Things Engineering*, *Jiangnan University*, b. *Key Laboratory of Advanced Process Control for Light Industry* (*Ministry of Education*), *Jiangnan University*, *Wuxi Jiangsu* 214122, *China*)

Abstract: Industrial processes often contain significant strong nonlinearity and time-varying behavior. Traditional extreme learning machine (ELM) based soft sensor sometimes fails to make use of data information effectively and has poor prediction performance. This paper proposed an improved particle swarm optimization algorithm. It has better convergence speed and search ability than standard particle swarm optimization (PSO) . This algorithm with an adaptive weight updates learning factor linearly and plays an important role in finding the optimal penalty coefficient and kernel parameter of ELM. The butane concentration of debutanizer column process simulation results verify that the proposed method has good prediction accuracy and generalization performance.

Key words: soft sensor modeling; extreme learning machine; particle swarm optimization; adaptive weight

0　引言

实际的工业生产过程中，存在许多与产品质量密切相关的质量变量，由于经济或技术局限难以进行在线实时监测和控制，如醋酸共沸精馏中的醋酸浓度[1]、油品黏度及组分、青霉素发酵过程产物浓度[2]等，软测量技术通过构建易测辅助变量和难测主导变量之间的数学模型，可以很好地解决这个问题[3]。常用的软测量模型如主元回归(principal component regression, PCR)[4]、偏最小二乘(partial least squares, PLS)[5]、人工神经网络(artificial neural networks, ANN)[6]、支持向量机(support vector machines, SVM)[7]、高斯过程回归(Gaussian process regression, GPR)[8]等都可以得到良好的预测结果。

软测量建模方法中，极限学习机(extreme learning machine, ELM)[9]由于具有预测精度高、学习速度快、参数少等优点，近几年被广泛应用于工业过程。相比于传统神经网络[10]、最小二乘支持向量机[11]等，ELM 模型克服了梯度下降学习算法需要多次迭代的缺点，能在满足所需精度要求的同时，提高神经网络的学习速度。然而 ELM 也存在一些自身的不足，如很难精确找到最优超参数、其泛化性能和预测精度也有待进一步提高等。尤其是参数选取不当，会导致网络学习性能严重不足，对查询样本的预测能力变差，从而不能满足实际工业的要求。目前，对于惩罚系数和核宽两个参数常采用人工选取的方式，该方式建立在大量的实验基础之上，花费的时间较长，且参数设置不当也会导致模型精度不高，因此，一些学者将群智能优化算法应用于 ELM 的惩罚系数和核宽的选取。Wang 等人[12]提出用灰狼优化(grey wolf optimization, GWO) 算法选取极限学习机的惩罚系数和核宽，所建模型具有更高的分类精度。Lyu 等人[13]将改进的细菌觅食优化(improved bacterial foraging optimization, IBFO) 算法应用于 ELM 超参数的选取，且在躯体化障碍患者的诊断方面取得了成功的应用。然而 GWO 算法存在收敛速度慢、容易陷入局部最优的缺点，IBFO 算法需要较多的参数设置，且寻优性能很大程度上受到趋化步长取值大小的影响，若优化算法的参数选取不当，仍然无法达到提高极限学习机模型泛化性能的目的。粒子群优化(particle swarm optimization, PSO) [14]最早是由 Kennedy 和 Eberhart 在 1995 年提出的，该算法原理清晰、可调参数少且收敛速度快，在机器学习算法的性能提升方面得到了较大的应用。Lin 等人[15]利用 PSO 算法对 SVM 进行特征选择以及最优超参数选取，明显提高 SVM 的泛化能力和分类准确性。陶新民等人[16]提出了一种改进的粒子群-K 均值聚类算法，该算法充分利用 PSO 全局搜索能力强的特性弥补了 K-均值聚类算法易陷入局部最优

收稿日期：2018-11-07；**修回日期：**2019-01-02　　**基金项目：**国家自然科学基金资助项目（61773182）；江苏高校优势学科建设工程资助项目（PAPD）

作者简介：盛晓晨（1996-），男，浙江金华人，硕士研究生，主要研究方向为复杂工业过程建模；熊伟丽（1978-），女，江苏无锡人，教授，博导，博士，主要研究方向为复杂工业过程建模与优化、智能软测量技术（greenpre@163.com）；史旭东（1993-），男，江苏常州人，硕士研究生，主要研究方向为复杂工业建模.

的缺陷。Grimaldi 等人[17]将 PSO 算法应用于神经网络权值配置，极大地改善了网络的学习性能。标准 PSO 算法在处理一些简单函数优化、参数选取等问题时体现出良好的寻优性能，但是在求解大量非线性、多峰函数的优化问题时存在早熟收敛、精度不高、稳定性差等缺陷，有时不容易得到全局最优解。

本文通过引入一种基于高斯下降的惯性权重改变策略，并线性变化学习因子，将 PSO 算法进行改进，以更快地跳出局部最优，提高网络收敛速度，加快对全局最优解的搜索，寻优效果得到显著的提升；并将改进的 PSO 算法用于选取 ELM 的惩罚系数和核宽，得到由最优超参数确定的 ELM 网络结构，不仅克服了由于参数选取不当导致神经网络收敛性差的缺点，而且明显提高了 ELM 模型的泛化能力和预测精度。通过脱丁烷塔过程的仿真研究，并通过不同建模方法的比较，对所提方法的有效性和精度进行了验证。

1　极限学习机

极限学习机是 2004 年由 Huang 等人提出的一种基于单隐层前馈神经网络(single-hidden layer feedforward networks, SLFNs) 的学习算法[18]，该算法随机给定输入层权值和隐含层节点偏置，只需设置网络的隐含层节点个数，便可通过求解线性方程组的最小二乘解获得输出权值，以此得到最优解[9]。

对于含有 N 个不同样本的训练集 $\{\boldsymbol{X},\boldsymbol{Y}\}=\{\boldsymbol{x}_i,\boldsymbol{y}_i\}_{i=1}^{N}$，其中 $\boldsymbol{x}_i=[x_{i1},x_{i2},\cdots,x_{in}]^{\mathrm{T}}\in R^n$，$\boldsymbol{y}_i=[y_{i1},y_{i2},\cdots,y_{im}]^{\mathrm{T}}\in R^m$，$n$ 为输入样本的维数，m 为输出样本的维数，该神经网络的输出可表示为

$$f(\boldsymbol{x}_i)=\boldsymbol{h}(\boldsymbol{x}_i)\boldsymbol{\beta}\ ,\ i=1,2,\cdots,N \tag{1}$$

其中：$\boldsymbol{h}(\boldsymbol{x}_i)$ 表示隐含层相对于 $\boldsymbol{x}_i$ 的输出向量，即隐含层输出矩阵，$\boldsymbol{\beta}$ 表示输出权重。

ELM 的训练目标是最小化训练误差和输出权值范数，可表示为一个带约束的优化问题：

$$\begin{cases}\min\limits_{\boldsymbol{\beta}\in R^{n_h\times m}} \dfrac{1}{2}\|\boldsymbol{\beta}\|^2+\dfrac{C}{2}\sum\limits_{i=1}^{N}\|\boldsymbol{e}_i\|^2\\ s.t.\boldsymbol{h}(\boldsymbol{x}_i)\boldsymbol{\beta}=\boldsymbol{y}_i^{\mathrm{T}}-\boldsymbol{e}_i^{\mathrm{T}},i=1,2,\cdots,N\end{cases} \tag{2}$$

其中：$\boldsymbol{e}_i\in \boldsymbol{R}^m$ 是第 i 个训练样本的训练误差；C 是惩罚系数。求解该优化问题，可得到

$$\boldsymbol{\beta}=\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\frac{\boldsymbol{I}_N}{C}\right)^{-1}\boldsymbol{Y} \tag{3}$$

其中：$\boldsymbol{H}=\left[\boldsymbol{h}(\boldsymbol{x}_1)^{\mathrm{T}},\boldsymbol{h}(\boldsymbol{x}_2)^{\mathrm{T}},\cdots,\boldsymbol{h}(\boldsymbol{x}_N)^{\mathrm{T}}\right]^{\mathrm{T}}$；$\boldsymbol{Y}=[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_N]^{\mathrm{T}}$；$\boldsymbol{I}_N$ 是维数为 N 的单位矩阵。应用 Mercer’s 条件定义核矩阵[19]：

$$\begin{cases}\boldsymbol{\Omega}_{ELM}=\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\\ \Omega_{ELMi,j}=h(\boldsymbol{x}_i)h(\boldsymbol{x}_j)=K(\boldsymbol{x}_i,\boldsymbol{x}_j)\end{cases} \tag{4}$$

在本文中选用高斯核函数：

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\exp\left\{-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{2\sigma^2}\right\} \tag{5}$$

其中：σ 为核函数的宽度参数，决定了核函数的径向作用范围。此时，ELM 的输出可由式(6) 表示。

$$\begin{aligned}f(x)&=\boldsymbol{h}(x)\boldsymbol{\beta}=\boldsymbol{h}(x)\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\frac{\boldsymbol{I}_N}{C}\right)^{-1}\boldsymbol{Y}\\ &=\begin{bmatrix}K(\boldsymbol{x},\boldsymbol{x}_1)\\ \vdots\\ K(\boldsymbol{x},\boldsymbol{x}_N)\end{bmatrix}^{\mathrm{T}}\left(\boldsymbol{\Omega}_{ELM}+\frac{\boldsymbol{I}_N}{C}\right)^{-1}\boldsymbol{Y}\end{aligned} \tag{6}$$

2　改进的粒子群优化算法

2.1　标准 PSO 算法

在 D 维空间内，种群规模为 N 的粒子群可表示为 $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N)^{\mathrm{T}}\in \mathrm{R}^{N\times D}$，粒子的位置和速度可分别表示为 $\boldsymbol{x}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{iD})^{\mathrm{T}}$ 和 $\boldsymbol{v}_i=(\boldsymbol{v}_{i1},\boldsymbol{v}_{i2},\cdots,\boldsymbol{v}_{iD})^{\mathrm{T}}$，粒子和整个种群在经迭代后得到的历史最好位置，可分别表示为个体极值 $p_{\mathrm{best}}=(p_{i1},p_{i2},\cdots,p_{iD})^{\mathrm{T}}$ 和群体极值 $g_{\mathrm{best}}=(g_{i1},g_{i2},\cdots,g_{iD})^{\mathrm{T}}$。

在寻找最优值的迭代中，每个粒子 $\boldsymbol{x}_i$ 通过个体极值和群体极值来调整自身的位置和速度，即

$$v_{id}^{k+1}=v_{id}^{k}+c_1r_1\left(p_{id}^{k}-x_{id}^{k}\right)+c_2r_2\left(p_{gd}^{k}-x_{id}^{k}\right) \tag{7}$$

$$x_{id}^{k+1}=x_{id}^{k}+v_{id}^{k} \tag{8}$$

其中：$i=1,2,\cdots,N$；$d=1,2,\cdots,D$；k 为当前迭代次数；$c_1,c_2\geq 0$ 为学习因子；r_1,r_2 均为介于 (0,1) 之间的随机数。

为了调节粒子自身历史速度对当前时刻的影响，以达到更好的优化效果，Shi 等人[14]在式 (7) 中引入了惯性权重 ω，于是得到

$$v_{id}^{k+1}=\omega v_{id}^{k}+c_1r_1\left(p_{id}^{k}-x_{id}^{k}\right)+c_2r_2\left(p_{gd}^{k}-x_{id}^{k}\right) \tag{9}$$

式 (7) 和 (9) 称为标准粒子群优化(SPSO)算法。惯性权重 ω 主要调节局部搜索能力和全局搜索能力的平衡，一般取 $\omega=0.8$。学习因子 c_1,c_2 主要调节个体极值和群体极值对粒子的影响，通常取 $c_1=c_2=2$。

2.2　改进的 PSO

标准 PSO 算法在搜寻最优解时表现了良好的优化效果，但同时也存在一些不足，例如标准算法在容易过早陷入局部最优解[20]，影响全局最优解的搜索。为了解决这个问题，国内外学者在研究惯性权重和学习因子的基础上提出了许多改进的 PSO 算法。

惯性权重是粒子群算法中的一个重要参数，较大的惯性权重可以提高 PSO 的全局搜索能力，反之有助于算法的局部搜索。Shi[14]提出一种线性递减权重的 PSO 算法 (LDWPSO)，其惯性权重的增加而线性减小。Amoshahy 等人[21]提出了柔性指数非线性权重改变策略。Liang 等人[22]提出一种新的动态非线性变惯量权重自适应变异粒子群优化算法。文献[23]中介绍了一种新的惯性权重设置方式，在线性递减惯性权重的基础上做了改进，得到一种非线性递减权重的 PSO 算法 (记做 w_1PSO)。

相比于 SPSO 算法，LDWPSO 算法在初期具有更好的全局搜索能力，后期具有更好的局部搜索能力，但是该算法在迭代过程中全局搜索和局部搜索的比例不变，在求解复杂非线性问题时效果往往不理想。w_1PSO 算法在寻优性能方面有了一定的提升，算法的稳定性和精度也得到了提高。为了进一步改善 PSO 的寻优效果，使算法在搜索前期具有较快的收敛速度，同时也避免搜索后期过早陷入局部最优，本文受高斯函数正态分布特点的启发，引入一种高斯下降的惯性权重 ω 改变策略，并线性变化学习因子 c_1 和 c_2 [24]，以此得到改进的 PSO 算法 (记做 IPSO)，设置如下：

$$\omega=1.1\exp\left(-\frac{k^2}{(d\times k_{\max})^2}\right) \tag{10}$$

$$\begin{cases}c_1=c_{1\min}+(c_{1\max}-c_{1\min})\times\dfrac{k_{\max}-k}{k_{\max}}\\ c_2=c_{2\min}+(c_{2\max}-c_{2\min})\times\dfrac{k}{k_{\max}}\end{cases} \tag{11}$$

其中：$c_1\in[c_{1\min},c_{1\max}]$；$c_2\in[c_{2\min},c_{2\max}]$；$d$ 为常数，本文取 $d=0.8$。IPSO 算法通过采用动态递减惯性权重，并同时线性改变学习

因子的策略，平衡了算法的全局搜索和局部开发能力，保证了种群的多样性，在提高收敛速度的同时能够获得较好的优化结果。

图 1 给出了三种不同 PSO 算法的惯性权重变化曲线。IPSO 算法的惯性权重 ω 随迭代次数的增加呈现高斯递减的下降趋势，相比于其他两种算法，该算法在迭代初期具有较大的惯性权重，收敛速度较快，具有较强的全局搜索能力，能够进行大范围搜索；在迭代后期搜索到最优值大致范围时设置较小的惯性权重，偏重发挥算法的局部开发能力，完成精细搜索。而 IPSO 的学习因子 c_1 随迭代次数的增加而减小，c_2 随迭代次数的增加而增加。采用这样的方式也是为了使粒子的飞行状态在搜索初期主要受个体最优位置的影响，提升粒子在整个搜索空间的开发能力；而在搜索后期粒子的状态主要受邻近粒子最优位置的影响，有助于算法的局部开发。

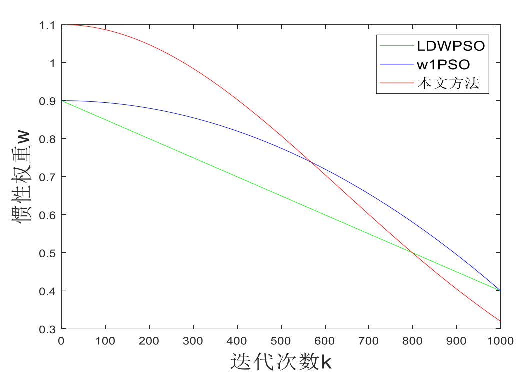


图 1 不同 PSO 算法的惯性权重变化曲线

Fig. 1 Inertia weight trend of different PSO algorithms

为了测试 IPSO 算法求解复杂问题的性能，选取标准测试函数中的 Rastrigin 函数对 SPSO、LDWPSO、w_1PSO、IPSO 四种算法进行仿真。该函数各变量之间无关联，且局部最优值随着正弦波动，是一个典型的具有大量局部最小值的复杂多峰函数[25]，函数表达式如式（12）所示。

$$f(x)=\sum_{i=1}^{n}[x_i^2-10\cos(2\pi x_i)+10] \tag{12}$$

其中：$x=[x_1,x_2,\cdots,x_n]^T$。现求取 Rastrigin 函数在二维空间 $[-5,5]\times[-5,5]$ 区域内的最小值，算法参数设置如下：种群规模设为 20，维数设为 2，$c_{1\max}=c_{2\max}=2.5$，$c_{1\min}=c_{2\min}=0.5$，$d=0.8$，$v\in[-0.02,0.02]$，$v_d\in[-5,5;-5,5]$。最大迭代次数设为 200 次，算法各实验 50 次，测试结果如表 1 所示。

表 1 不同算法的测试结果比较

Table 1 Test results comparison of different algorithms

方法	最差值	最优值	平均值	方差	标准差
SPSO	3.5746	3.53×10^{-6}	0.3362	0.7312	0.8700
LDWPSO	2.5912	1.59×10^{-9}	0.3085	0.4379	0.6618
w_1PSO	2.5800	1.46×10^{-9}	1.74×10^{-1}	0.2465	0.4965
IPSO	0.2404	4.03×10^{-12}	6.34×10^{-3}	0.0012	0.0339

从表 1 可以看出，相比于其他三种算法，本文提出的 IPSO 算法经多次迭代后能够以较高概率收敛于全局最优解，寻优效果更好，可利用该算法优化 ELM 的惩罚系数和核宽，以期得到预测精度更高的训练模型，改善回归性能。

3 基于改进粒子群优化的极限学习机

ELM 的泛化性能在很大程度上取决于其参数包括惩罚系数和核宽的合理选取。因此，本文过改进的粒子群算法对极限学习机的超参数进行寻优，以获得精度更高的 ELM 预测模型。定义“粒子”为寻优解空间中的可能解 (C,σ)，选取式(13)所示的均方根误差（root mean square error, RMSE）作为粒子群算法的适应度函数[26]，将计算所得的适应度值记为 $\text{fitness}(i)$。

$$\text{RMSE}=\sqrt{\frac{1}{n_{\text{test}}}\sum(y_{\text{test}}-y_{\text{pred}})^2} \tag{13}$$

其中：n_{test} 为测试样本个数；y_{test} 为测试样本真实值；y_{pred} 为模型预测输出值。

IPSO 优化 ELM 参数的训练步骤如下：

a)初始化粒子群算法相关参数，主要有种群规模 N，空间维数 D，学习因子 c_1 和 c_2，惯性权重 ω，迭代次数 $k_{\max}$，粒子的速度限制 $[v_{\min},v_{\max}]$ 和位置限制 $[x_{d\min},x_{d\max}]$ 等，计算第 1 个粒子 x_1 的适应度值 $\text{fitness}(1)$，令最优适应度值 $\text{fitness(best)}=\text{fitness}(1)$，并将粒子 x_1 作为 p_{best}。

b)随机初始化输入权值 ω_i 和隐含层偏置 b_i，构建 ELM 神经网络结构，并随机产生一组数据 (C,σ) 作为初始解空间。

c)采用粒子 (C,σ) 对应的优化后的 ELM 模型对样本进行训练，计算适应度值并评价。每个粒子 (C,σ) 通过比较个体极值的适应度值和全局极值的适应度值来调整自身的位置和速度。

d)判断是否达到迭代次数，若达到，则算法终止，并给出搜索结果，否则继续迭代得到新的粒子群。

e)将最终优化得到的粒子 (C,σ) 作为极限学习机的惩罚系数和核宽，以此建立优化的 ELM 模型。

IPSO 优化 ELM 参数的算法流程如图 2 所示。

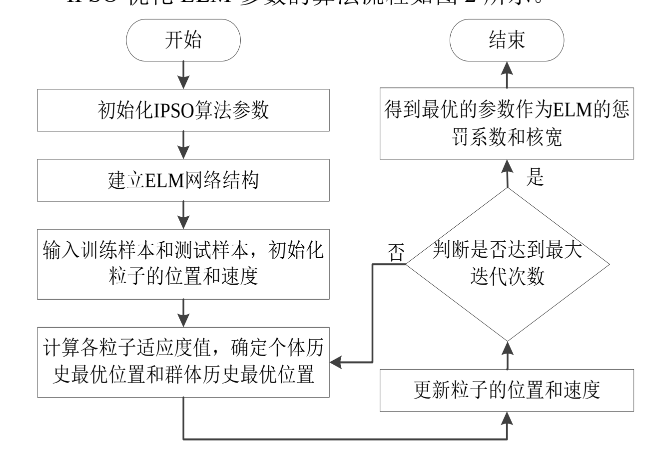


图 2 IPSO 优化 ELM 参数的算法流程

Fig. 2 Flow chart of ELM algorithm with parameters optimized by IPSO

4 仿真研究

4.1 ELM 算例仿真

为了测试 ELM 的惩罚系数和核宽对其学习性能的影响，此处对函数 $f(x)=x\sin x$ 作回归分析。在本次实验中，随机生成 200 组数据 $\{x_i,f(x_i)+\zeta_i\}$，其中 100 组作为训练样本，100 组作为测试样本，ζ_i 为一组噪声，服从正态分布 $\zeta_i\sim N(0,0.05)$。选取 RMSE 和跟踪系数（tracking precision, TP）[27]作为评价 ELM 拟合效果的性能指标。TP 具体定义如式（14）所示。

$$\text{TP}=1-\frac{\sum(y_{test}-y_{pred})^2}{\sum y_{test}^2} \tag{14}$$

考虑到测试结果的波动，每次实验做了 10 组并求取平均值，分别测试 ELM 的 RMSE 和 TP，所得数据如表 2 和 3 所示。

表 2　惩罚系数和核宽对 RMSE 的影响

Table 2　Effect of different penalty coefficient and kernel parameter on RMSE

C σ	0.01	0.5	5	50	100	**200**	1000
0.01	0.3235	0.0720	0.0476	0.0559	0.0558	0.0540	0.0514
0.05	0.2678	0.0700	0.0561	0.0531	0.0532	0.0553	0.0575
0.1	0.3277	0.0567	0.0553	0.0569	0.0481	0.0482	0.0525
0.3	0.3115	0.0728	0.0574	0.0514	0.0488	0.0470	0.0570
0.5	0.2604	0.0829	0.0505	0.0501	0.0532	**0.0459**	0.0469
1	0.2589	0.0726	0.0537	0.0463	0.0463	0.0525	0.0508
5	0.2829	0.1414	0.0759	0.0560	0.0474	0.0545	0.0506
50	0.3493	0.2302	0.1032	0.0685	0.0688	0.0736	0.0693
100	0.2695	0.2606	0.1595	0.0657	0.0855	0.0724	0.0690

表 3　惩罚系数和核宽对 TP 的影响

Table 3　Effect of different penalty coefficient and kernel parameter on TP

C σ	0.01	0.5	5	50	100	**200**	1000
0.01	0.2402	0.9484	0.9639	0.9541	0.9560	0.9604	0.9547
0.05	0.3774	0.9371	0.9545	0.9555	0.9594	0.9584	0.9603
0.1	0.2955	0.9591	0.9433	0.9564	0.9701	0.9614	0.9608
0.3	0.3054	0.9277	0.9573	0.9667	0.9690	0.9698	0.9501
0.5	0.3236	0.8997	0.9514	0.9687	0.9589	**0.9709**	0.9670
1	0.1984	0.9265	0.9573	0.9693	0.9677	0.9620	0.9625
5	0.0580	0.7397	0.9187	0.9590	0.9615	0.9547	0.9680
50	0.0045	0.2428	0.8272	0.9199	0.9379	0.9193	0.9303
100	0.0028	0.1196	0.6204	0.9192	0.9059	0.9282	0.9296

从表 2 和 3 中可以看出，ELM 的泛化性能与选取的惩罚系数和核宽大小有很大关系，在本测试中，选取 $C=200$，$\sigma=0.5$ 时 ELM 对该函数的拟合性能较好。在参数选取合适时，ELM 体现了良好的学习能力，在噪声存在的情况下仍能以较高精度完成对函数值的预测；反之，ELM 泛化性能较差。

4.2　脱丁烷塔过程仿真

脱丁烷塔过程是化工过程石油炼制中脱硫和石脑油分离装置的重要组成部分[28]，塔底丁烷浓度对石油炼制影响较大且难以直接采用仪器测量，需要对该工业过程进行软测量建模。选取 x_1 顶层温度、x_2 顶层压力、x_3 回流流量、x_4 进入下一流程流量、x_5 第 6 塔板温度、x_6 塔顶温度 1、x_7 塔顶温度 2 等七个可测变量作为辅助变量，主导变量为塔底丁烷浓度。图 3 给出了脱丁烷塔过程示意图，并标注了辅助变量的测量点位置。

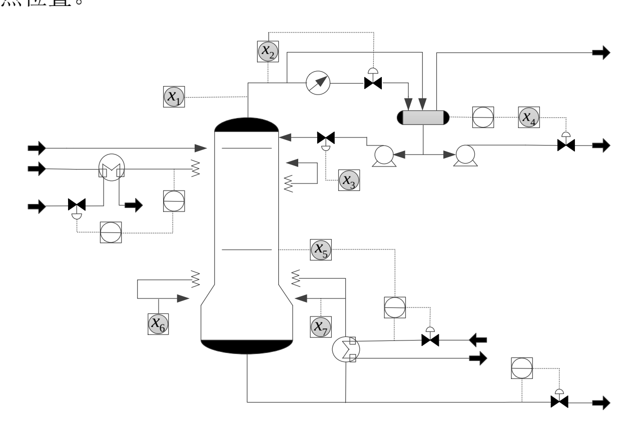


图 3　脱丁烷塔过程示意图

Fig. 3　Schematic diagram of debutanizer process

共收集 2 000 组脱丁烷塔过程数据，其中 1 000 组作为训练样本集，1 000 组作为测试样本集，建立三种不同的软测量模型：ELM 模型、SPSO 优化的 ELM 模型（记为 PSOELM 模型）、IPSO 优化的 ELM 模型（简记为 IELM 模型），分别进行仿真实验并比较。种群规模设为 100，最大迭代次数设为 100 次，$v\in[-0.5,0.5]$，$x_d\in[6,3;2,-2]$，其余参数设置同章节 2.2。选取 RMSE［式 (13)］、TP［式 (14)］和相关系数 (Correlation Coefficient, COR)［式 (15)］作为预测性能指标。

$$COR=\frac{\mathrm{Cov}(y_{\mathrm{test}},y_{\mathrm{pred}})}{\sqrt{\mathrm{var}(y_{\mathrm{test}})}\cdot\sqrt{\mathrm{var}(y_{\mathrm{pred}})}} \tag{15}$$

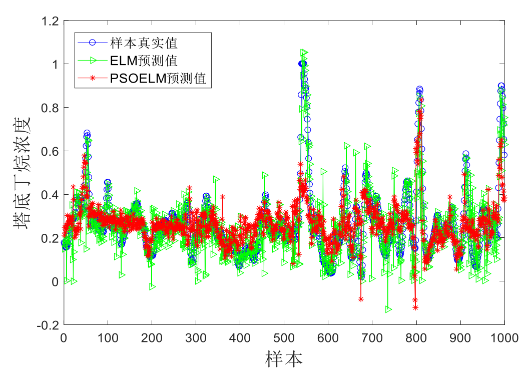
表 4 给出了三种建模方法对丁烷浓度的预测结果比较。从表中可以看出，经 IPSO 算法优化后得到的 IELM 模型对塔底丁烷浓度的预测精度明显高于传统 ELM 模型，进一步与 PSOELM 模型相比，IELM 模型对丁烷浓度的预测误差减小，在泛化性能方面也有了一定的提升，表明本文方法可以有效提高模型的学习能力和预测精度，能够较好地完成对丁烷浓度的预测。

表 4　三种建模方法的预测结果比较

Table 4　Prediction results comparison of three modeling methods

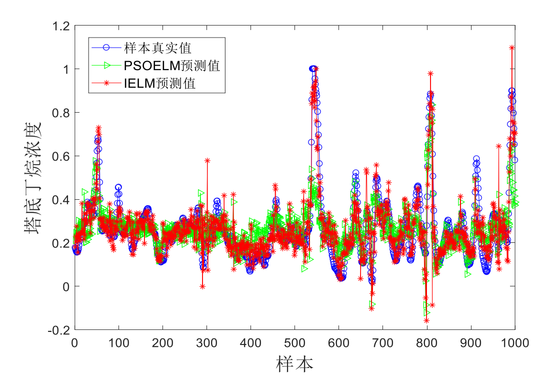
方法	RMSE	COR	TP
ELM	0.1051	0.7301	0.5323
PSOELM	0.0803	0.8462	0.7146
IELM	**0.0746**	**0.8830**	**0.7643**

为了更直观地显示 IELM 算法的学习效果，图 4 给出了三种建模方法得到的软测量模型对丁烷浓度的预测结果。由图可知，IELM 模型对丁烷浓度的预测值曲线更紧凑得跟随于真实值曲线，预测误差在 0 附近波动幅度更小，说明 IELM 模型输出的丁烷浓度预测值更接近真实值，预测结果与真实情况的更加吻合，预测产生的误差较小。相比于传统 ELM 模型仿真结果，本文方法对丁烷浓度的拟合能力更强，所建模型的泛化能力更强，预测精度更高。



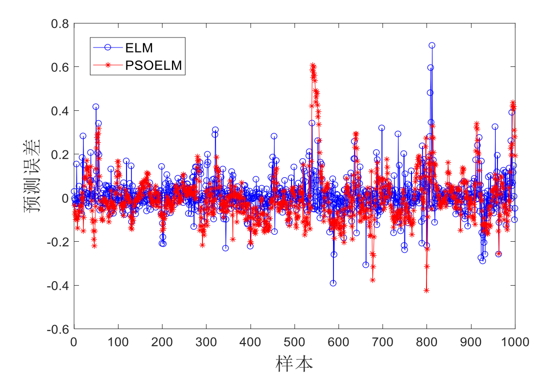
(a) ELM、PSOELM 对丁烷浓度的预测结果

(a) Prediction results of ELM and PSOELM for C4 concentration



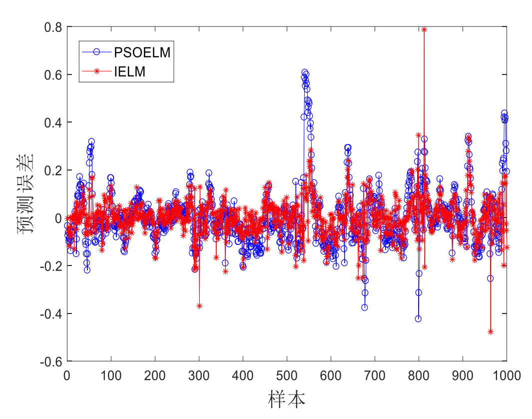
(b) PSOELM、IELM 对丁烷浓度的预测结果

(b) Prediction results of PSOELM and IELM for C4 concentration



(c) ELM、PSOELM 对丁烷浓度的预测误差

(c) Prediction errors of ELM and PSOELM for C4 concentration



(d) PSOELM、IELM 对丁烷浓度的预测误差

(d) Prediction errors of PSOELM and IELM for C4 concentration

图 4　ELM、PSOELM 和 IELM 的仿真结果

Fig. 4　Simulation results of ELM、PSOELM and IELM

为了进一步说明本文所提算法的有效性，图 5 给出了三种建模方法对丁烷浓度预测结果的散点图。可以看出，IELM 模型仿真得到的散点更加接近于黑色对角线，而传统 ELM 模型和 PSOELM 模型得到的散点分布则相对较为分散，说明本文方法建立的模型得到的丁烷浓度预测结果更接近真实情况，预测精度明显得到提高。

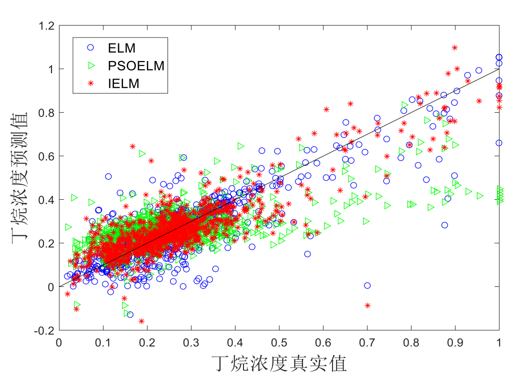


图 5　ELM、PSOELM 和 IELM 预测结果散点图

Fig. 5　Scatter plot of ELM, PSOELM and IELM prediction results

5　结束语

利用一种改进的粒子群优化算法对极限学习机的惩罚系数和核宽两个参数同时进行优化，得到性能更优的极限学习机软测量模型，在增强模型泛化能力的同时，提高了模型对非线性工业过程质量变量的预测精度。通过脱丁烷塔过程的应用仿真，验证了该方法的有效性和精度，对于复杂工业过程的软测量建模具有一定的参考价值。

参考文献：

[1] 潘红芳. 极限学习机的研究及其在醋酸精馏软测量建模中的应用 [D]. 华东理工大学, 2014. (Pan Hongfang. Research on extreme learning machine and its application in acetic acid soft sensor modeling [D]. East China University of Science and Technology, 2014.)

[2] 熊伟丽, 姚乐, 徐保国. 基于 EM 算法的青霉素发酵过程多阶段融合建模 [J]. 化工学报, 2014, 65 (12): 4935-4941. (Xiong Weili, Yao Le, Xu Baoguo. Multi-stage fusion modeling of penicillin fermentation process based on EM algorithm [J]. CIESC Journal, 2014, 65 (12): 4935-4941.)

[3] 曹鹏飞, 罗雄麟. 化工过程软测量建模方法研究进展 [J]. 化工学报, 2013, 64 (3): 788-800. (Cao Pengfei, Luo Xionglin. Modeling of soft sensor for chemical process [J]. CIESC Journal, 2013, 64 (3): 788-800.)

[4] Niu Zhiguang, Wang Chong, Zhang Ying, et al. Leakage rate model of urban water supply networks using principal component regression analysis [J]. Journal of Tianjin University, 2018, 24 (2): 172-181.

[5] Zheng Junhua, Song Zhihuan. Semisupervised learning for probabilistic partial least squares regression model and soft sensor application [J]. Journal of Process Control, 2018, 64: 123-131.

[6] Willis M J, Montague G A, Massimo C D, *et al.* Artificial neural networks in process estimation and control [J]. Automatica, 1992, 28 (6): 1181-1187.

[7] Cortes C. Support vector network [J]. Machine Learning, 1995, 20 (3): 273-297.

[8] Xiong Weili, Shi Xudong. Soft sensor modeling with a selective updating strategy for Gaussian process regression based on probabilistic principle component analysis [J]. Journal of the Franklin Institute, 2018, 35: 5336-5349.

[9] Huang Guangbin, Zhu Qinyu, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks [C]// Proc of IEEE International Joint Conference on Neural Networks,.2005: 985-990.

[10] Sun Kai, Liu Jialin, Kang Jialin, *et al.* Development of a variable selection method for soft sensor using artificial neural network and nonnegative garrote [J]. Journal of Process Control, 2014, 24 (7): 1068-1075.

[11] 赵彦涛, 单泽宇, 常跃进, 等. 基于 MI-LSSVM 的水泥生料细度软测量建模 [J]. 仪器仪表学报, 2017 (2): 487-496. (Zhao Yantao, Shan Zeyu, Chang Yuejin, *et al.* Soft sensor modeling for cement fineness based on least squares support vector machine and mutual information [J]. Chinese Journal of Scientific Instrument, 2017 (2): 487-496.)

[12] Wang Mingjing, Chen Huiling, Li Huaizhong, *et al.* Grey wolf optimization evolving kernel extreme learning machine: application to bankruptcy prediction [J]. Engineering Applications of Artificial Intelligence, 2017, 63: 54-68.

[13] Lyu Xinen, Chen Huiling, Zhang Qian, *et al.* An improved bacterial-foraging, optimization-based machine learning framework for predicting the severity of somatization disorder [J]. Algorithms, 2018, 11 (2): 17.

[14] Shi Yuhui, Eberhart R. Modified particle swarm optimizer [C]// Proc. of IEEE ICEC Conference. 1999: 69-73.

[15] Lin S W, Ying K C, Chen S C, *et al.* Particle swarm optimization for parameter determination and feature selection of support vector machines [J]. Expert Systems with Applications, 2008, 35 (4):

1817-1824.

[16] 陶新民，徐晶，杨立标，等. 一种改进的粒子群和K均值混合聚类算法 [J]. 电子与信息学报, 2010, 32 (1): 92-97. (Tao Xinmin, Xu Jing, Yang Libiao, *et al.* Improved cluster algorithm based on K-means and particle swarm optimization [J]. Journal of Electronics & Information Technology, 2010, 32 (1): 92-97.)

[17] Grimaldi E A, Grimaccia F, Mussetta M, *et al.* PSO as an effective learning algorithm for neural network applications [C]// Proc of International Conference on Computational Electromagnetics and ITS Applications. 2004: 557-560.

[18] Huang Guangbin, Song Shiji, Gupta J N, *et al.* Semi-supervised and unsupervised extreme learning machines [J]. IEEE Trans on Cybernetics, 2014, 44 (12): 2405-2417.

[19] 马超，徐守祥，刘远东. 集成优化核极限学习机的冠心病无创性诊断 [J]. 计算机应用研究, 2017, 34 (6): 1671-1676. (Ma Chao, Xu Shouxiang, Liu Yuandong. Optimized kernel extreme learning machine based on ensemble method for heart diseases [J]. Application Research of Computers, 2017, 34 (6): 1671-1676.)

[20] 王存睿，段晓东，刘向东，等. 改进的基本粒子群优化算法 [J]. 计算机工程, 2004, 30 (21): 35-37. (Wang Cunrui, Duan Xiaodong, Liu Xiangdong, *et al.* A modified basic particle swarm optimization algorithm [J]. Computer Engineering, 2004, 30 (21): 35-37.)

[21] Amoshahy M J, Shamsi M, Sedaaghi M H. A novel flexible inertia weight particle swarm optimization algorithm [J]. Plos One, 2016, 11 (8): e0161558.

[22] Liang Hongtao, Kang Fengju. Adaptive mutation particle swarm algorithm with dynamic nonlinear changed inertia weight [J]. Optik-International Journal for Light and Electron Optics, 2016, 127 (19): 8036-8042.

[23] 史峰，王辉，郁磊，等. MATLAB智能算法30个案例分析 [M]. 北京: 北京航空航天大学出版社, 2011. (Shi Feng, Wang Hui, Yu Lei, *et al.* 30 case analysis of MATLAB intelligent algorithm [M]. Beijing: Beihang University Press, 2011.)

[24] Ratnaweera A, Halgamuge S K, Watson H C. Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients [J]. IEEE Trans on Evolutionary Computation, 2004, 8 (3): 240-255.

[25] 鲁姝颖. 粒子群优化算法的几种改进算法及应用 [D]. 徐州：中国矿业大学, 2014. (Lu Shuying. Some improvements and applications of particle swarm optimization algorithm [D]. Xuzhou: China University of Mining and Technology, 2014.)

[26] 王杰，毕浩洋. 一种基于粒子群优化的极限学习机 [J]. 郑州大学学报：理学版，2013，45 (1)：100-104. (Wang Jie, Bi Haoyang. An extreme learning machine based on particle swarm optimization [J]. Journal of Zhengzhou University: Engineering Science, 2013, 45 (1): 100-104.)

[27] Zhang Wei, Li Yanjun, Xiong Weili, *et al.* Adaptive soft sensor for online prediction based on enhanced moving window GPR [C]// Proc of IEEE International Conference on Control, Automation and Information Sciences . 2015: 291-296.

[28] Fortuna L, Graziani S, Xibilia M G. Soft sensors for product quality monitoring in debutanizer distillation columns [J]. Control Engineering Practice, 2005, 13 (4): 499-508.